REMARKS

Claims 44, 45, 46, 50, and 51 have been cancelled. Claims 1-43 and 47-49 are now pending in the application. Claims 3, 4, 11, 12, 13, 14, 17, 18, 19, 20, 22, 23, 30, 31, 32, 33, 36, 40, 41, 42, 43, 47, 48, and 49 have been amended. No new matter has been added by amendment. Reexamination and reconsideration of the claims as amended are respectfully requested.

Claim Objections

2. The objection to claims 1, 21, 37, and 40 were withdrawn in light of the claim amendments that deleted the blank and added the ATCC accession number. For the file record a copy of the ATCC deposit slip for PH51H is included in the current response.

Claim Rejections Under Double Patenting

8. The Examiner rejects claims 14, 17, 33, 36, 41, 43, 45, and 46 under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-26 of U.S. Patent No. 6,188,001 for the reasons of record. Applicant again traverses the rejection.

Claims 45 and 46 have been cancelled. Claims 17, 33, 36, 41, and 43 have been amended. Claims 36, 41, and 43 are now limited to PH51H-progeny that have been produced through one out cross. Claims 17 and 33 are limited to PH51H-progeny that are within a pedigree distance of two or less crosses from PH51H.

The Examiner states that, "The instantly claimed plants that are derived from crosses and breeding programs are not patentably distinct from the patented plants that are derived from crosses and breeding programs involving PH1W0, as they can express traits that are also expressed by the patented plants. " Applicant respectfully disagrees with the Examiner. Applicant submits that PH51H is clearly differentiated from PH1W0. One would not be able to obtain PH51H through the modification of the maize inbred taught in patent 6,188,001 because PH51H comprises a unique and nonobvious combination of previously unknown and nonobvious genetics. Further, plants derived from PH51H are also clearly differentiated. The corn genome contains enormous complexity, and it is not possible that the claimed plants derived from PH51H could have been produced without the use of PH51H. In particular, PH1W0 could not be substituted as the starting material to produce the claimed plants derived from PH51H. The

genetics of PH51H comprise linkage groups and polymorphisms unique to PH51H, and it would be impossible to completely remove the contribution of PH51H to its progeny within one breeding cycle.

In light of the above, Applicant respectfully requests that the Examiner reconsider and withdraw the rejection to claims 1-49 due to double patenting or provide some clear evidence to establish why PH51H would be obvious over PH1W0. See In re Kaplan, 789 F. 2d 1580,229 U.S.P.Q. 683.

Claim Rejections – 35 USC § 112, second paragraph

9. The Examiner rejects claims 3, 5, 14, 22, 33, 40-46, 50, and 51 under 35 USC § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The Examiner rejects claims 3 and 22 because of the recitation of "wherein the plant is male sterile." The Applicant traverses the rejection but has amended the claims for clarification purposes and to expedite prosecution. Claims 50 and 51 have been cancelled. Claim 3 has been amended by replacing "male sterile" with "detasseled" as suggested by the Examiner. The Examiner suggests that a new claim be directed towards a method of producing a male sterile maize plant comprising transforming the maize plant of claim 2 with a nucleic acid that confers male sterility, and another claim directed towards the male sterile plant produced by the method of transforming. Claim 22 has been amended and now reads, "The maize plant of claim 2, wherein genes controlling male sterility have been transferred into said maize plant through crossing that utilizes PH51H as a recurrent parent and wherein said plant has essentially the same morphology and physiology of inbred line PH51H other than the trait of male sterility." Starting on page 1, line 34 and going through line 14 of page 3 of the specification it states that various genes, nuclear and cytoplasmic, have been used to control sterility in maize plants. In the specification on page 4, lines 7-19, it states, "Backcrossing can be used to transfer a specific desirable trait from one inbred or source to an inbred that lacks that trait. This can be accomplished, for example, by first crossing a superior inbred (recurrent parent) to a donor inbred (non-recurrent parent), that carries the appropriate gene(s) for the trait in question. The progeny of this cross is then mated back to the superior recurrent parent followed by selection in the resultant progeny for the desired trait to be transferred from the non-recurrent parent." A hybrid developed from inbreds containing the transferred gene(s) is essentially the same as a

hybrid developed from the same inbreds without the transferred gene(s)." The technique of backcrossing male sterility genes into an inbred maize plant is well known and well understood to one of ordinary skill in the art. The technique has been successfully used since the 1950's (see pages 585-586 of Wych, 1988, included in the information disclosure statement). The amendments contain no new matter. The Applicant requests reexamination and reconsideration of the claims as amended.

The Examiner rejects claims 5 and 24 because there is no antecedent basis for "protoplasts". The Applicant has amended claims 4 and 23. The claims now read, " A tissue culture of regenerable cells or protoplasts from the plant of claim 2 {21}." Thus the term "protoplasts" in claims 5 and 24 that depend from claims 4 and 23 respectively, has proper antecedent basis. The amendments place the claims in condition for allowance.

The Examiner rejects claim 40 as being indefinite because of the recitation of "comprising" and suggests the insertion of the terms "F1 hybrid" be inserted. Claim 40 has been amended to specify the first generation (F1) plant. Dependent claim 40 has also been amended to reflect the change in claim 40.

Claim Rejections – 35 USC § 112, first paragraph

10. The Examiner rejects claims 9-20 and 28-49 under 35 USC § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor, at the time of the application was filed, had possession of the claimed invention. Applicant traverses the rejection.

The Examiner rejects claims 9, 10, 28, and 29 that are directed to F1 hybrids produced with PH51H as a parent. Applicant notes that a claim to the F1 hybrid made with a deposited inbred was expressly acknowledged without reservation by the United States Supreme Court In *J.E.M. Ag. Supply, Inc. v. Pioneer Hi-Bred Int'l, Inc.*, 60 USPQ 2d 1865,1873 (S.Ct. 2001), when the Supreme Court wrote, "...a utility patent on an inbred plant line protects the line as well as all hybrids produced by crossing that inbred with another plant line."

Furthermore, one of ordinary skill in the art would know if they were using or one could easily identify if they were using PH51H. All F1 plants would have essentially the same genetic markers as the deposited PH51H. It is well known to anyone skilled in the art that a hybrid has a genome with one set of the alleles from each inbred. Therefore

the genetic profile exhibited in the deposit would be exhibited in the hybrid. As stated in the specification on page 15, lines 1-16, there are many laboratory-based techniques available for the analysis comparison and characterization of plant genotype such as Restriction Length Polymorphisms (RFLPs) and Simple Sequence Repeats (SSRs). Such techniques have been known for some time and may be used to identify whether or not PH51H was used to develop a hybrid. Applicant also submits to the Examiner the journal article by Berry et al. (2002). This article discusses the probability of identifying the parents of the hybrid by SSR data when neither parent is known. A copy of article by Berry et al. is attached to this Amendment and Request for Reconsideration as Appendix B. The results of the experiment showed that using 100 SSR loci markers resulted in correct parental ranking of inbreds for 53 out of 54 hybrids. Applicant also points out that any breeder of ordinary skill in the art will know the identity of both parents used to produce a hybrid.

The Examiner broadly rejects product claims encompassing any modification of PH51H, no matter how minor the modification or routine the modification is for a breeder of ordinary skill in the art to make.

As noted in the specification, the development of an inbred line is a time consuming and labor intensive activity. On average, between 10,000 to 20,000 lines are created and screened in order to develop any maize inbred line for which an Applicant files a patent application. Once developed, the inbred line is useful for two purposes: (1) to make commercial hybrids, and (2) as a source of breeding material for the development of new inbreds that retain its desired characteristics. A breeder desiring to make a line with similar traits to PH51H would be greatly advantaged by being able to use PH51H as starting material. This is because the linked genes arranged through Applicant's breeding efforts, and fixed in PH51H, can be maintained in the progeny of PH51H by a breeder of ordinary skill in the art. For example, if a breeder of ordinary skill in the art desired an early maturity version of PH51H, the breeder could cross PH51H to an earlier maturing variety, select for progeny with at least two desired PH51H traits that also express early maturity, and continue selecting for the traits of PH51H combined with early maturity. Optionally, the breeder could backcross to PH51H to obtain further genetic contribution from PH51H. The end result is the development of an inbred line with substantially all of the benefit of Applicant's work but with only a fraction of the effort.

Specifically, in rejecting the claims for lack of written description, the Examiner states, "the specification does not describe the plants produced by the corn breeding programs, transgenic PH51H plants, PH51H plants comprising single gene conversion(s), or by crosses wherein at least one ancestor is the corn variety PH51H, other than PH51H/PH1W2. The morphological and physiological traits of the corn plants that are crossed with PH51H, and with progeny of that cross, are unknown, and the description of progeny and descendents of corn plant PH51H are unknown. The description of corn plant PH51H is not indicative of the description of plants and seed produced by the breeding programs and crosses, or any of its descendents. The claimed invention also encompasses plants that express at least two 'PH51H traits' listed in claims 14, 33, and 46. However, to say that a plant expresses two traits of another plant is not sufficient information to describe that plant, as numerous corn plants express at least two of the same traits as those expressed by PH51H. Two plant traits do not provide any description of the other traits of a plant. It is possible that the claimed plants inherited the genes governing those traits from an ancestor other than plant PH51H. For, example, Piper (U.S. Patent No. 6,188,001) describes a corn plant, designated 'PH1W0,' which has at least two traits in common with PH51H, a relative maturity and adaptation to the Northcentral region traits, for example (col. 10, lines 50-62). The instantly claimed corn plants could have PH1W0 as an ancestor, as well as PH51H, in which the relative maturity rating and adaptation to the Northcentral region traits, for example, could have been inherited from PH1W0. The claims also encompass plants that do not have to express any of the traits that are expressed by PH51H."

Applicant notes that Examiner's comments a change of patent office policy. In numerous previous cases involving the protection of germplasm and progeny claims, including cases allowed after the recently adopted written description guidelines, the listing of traits was previously required by the patent office as a way to meet the written description requirement with respect to progeny. One reason for using traits as a means of description is because it was, and still is, technologically impossible to sequence the entire genome of a specific variety.

This situation is somewhat analagous to Ex Parte Tanksley, 37 USPQ2d. 1382. In that case the Examiner desired that Tanksley claim according to sequence data to "better characterize the cDNA clones" and "facilitate a complete search of the prior art" and issued a 112 first paragraph written description rejection. The Board held that "the section 112 rejection amounts to a requirement...that the appellants amend their claims

in a specified manner...We find no language in the statute or case law which would support that requirement." The Board, in treating the section 112 first paragraph rejection as a 112 second paragraph rejection, held that "In our judgement, a patent applicant is entitled to a reasonable degree of latitude in complying with the second paragraph of 35 U.S.C. 112 and the examiner may not dictate the literal terms of the claims for the stated purpose of facilitating a search of the prior art. Stated another way, a patent applicant must comply with 35 U.S.C. 112, second paragraph, but just how the applicant does so, within reason, is within applicant's discretion." Id. at 1386.

Applicant has amended claims 17 and 33 to limit the progeny covered to those within a pedigree distance of two crosses away from PH51H. Claims 36 and 41 are limited to one cross away from PH51H. Within the plant breeding arts breeders use pedigree as a means to characterize lines in reference to their progenitors. To those of ordinary skill in the art, this indicates that a line fewer crosses away from a starting line will be, as a whole, more highly related to the starting line. Thus, the work of the original breeder in developing the starting line will be retained in the closely related progeny. More specifically, traits and linkage groups present in PH51H will be retained in progeny that are within 2 outcrosses from PH51H. Applicant submits that characterization of the progeny of PH51H by virtue of their filial relationship is clearly within reason. Not only are filial descriptions used by breeders to evaluate materials for use in their breeding programs, but it is standard practice within the plant breeding industry for licensor's of inbred maize lines to retain a royalty from lines developed through the use of their inbreds. Those royalties are, in almost all cases, based on the filial relationship between the licensed inbred used in breeding and the progeny line commercialized. This provides evidence that those of ordinary skill in the art of plant breeding describe progeny in terms of pedigree.

Applicant also notes that the mere fact that the progeny have not been created does not prevent them from being patented. As stated in MPEP 2163 (3) (a), "An invention may be complete and ready for patenting before it has actually been reduced to practice." As stated in the written description guidelines "an applicant shows possession of the claimed invention by describing the claimed invention with all its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention. Possession may be shown in a variety of ways, including...by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention." 1255 Official

Gazette 140 (Feb. 5, 2002). Pedigree, which is a formula used by plant breeders, is a distinguishing identifying characteristic in compliance with the written description guidelines. Further, the Examiner must evaluate written description by the claimed invention with all of its limitations, including the limitation of being derived from PH51H.

PH51H-derived progeny are described by the fact that PH51H is utilized in a breeding program to make the PH51H-derived progeny, PH51H gives genetic contribution to the PH51H-derived progeny, and the genetics of PH51H are described by ATCC deposit of PH51H seed. By limiting the progeny to 2 or less crosses away from PH51H, the Examiner's concern that the progeny may be only distantly related to PH51H is addressed. In Enzo vs. Gen-Probe, U.S. State Court of Appeals for the Federal Circuit, 63 USPQ 2d 1609, the court reversed its prior decision regarding the insufficiency of the deposited genetic probes to meet the written description requirement. In so holding, the court stated, " As the deposited sequences are about 850, 8500, and 1300 nucleotides long, ..., there are at least hundreds of subsequences of the deposited sequences, an unknown number of which might also meet the claimed hybridization ratio. Moreover, Enzo's expert, Dr. Wetmur, stated that 'astronomical' numbers of mutated variations of the deposited sequence also fall within the scope of those claims, and that such broad claim scope is necessary to adequately protect Enzo's invention from copyists who could otherwise make minor change to the sequence and thereby avoid infringement while still exploiting the benefits of Enzo's invention. The defendants assert that such breadth is fatal to the adequacy of the written description. On the other hand, because the deposited sequences are described by virtue of a reference to their having been deposited, it may well be that various subsequences, mutations, and mixtures of those sequences are also described to one of skill in the art. We regard that question as an issue of fact...."

The issue of whether the progeny as now claimed satisfies the written description requirement is also an issue of fact. One of ordinary skill in the art would know if PH51H were utilized in a breeding program by looking at the breeding records and therefore would know if a progeny were derived from PH51H. PH51H is a unique inbred, as evidenced by the morphological and physiological traits given in Table 1, pages 17-19, of the application. Routinely used molecular techniques, discussed on page 15, lines 1-16 of the application, can be used to verify whether PH51H is within the pedigree of a line.

Applicant would also like to emphasize that PH51H cannot be derived through any other means then through PH51H seed and plant, nor can the influence of PH51H on the progeny be removed from a line within 2 outcrosses of PH51H. This fact also highlights the different perspective between the Examiner and the Applicant regarding the scope of the claims. The Examiner believes the claims to progeny to be of great breadth. However, to view these claims as being of great breadth merely because a large number of plants could theoretically fall within its scope ignores an essential limitation of the claim; that only a plant developed through the use of PH51H is within the scope of the claim. Such a plant could not be independently derived without the use of PH51H, so the claim would not in any way restrict the work of a breeder that did not in fact use PH51H. A breeder infringing such a claim must have made a conscious choice to use PH51H in order to obtain some or all of PH51H's desired characteristics. Compliance with the written description requirement is essentially a fact based inquiry that will "necessarily vary depending on the nature of the invention claimed." Vas-Cath v. Mahurkar, 935 F. 2d 1555 (citing In re DiLeone, 436 F2d. 1404, 1405). Thus, the compliance with the written description requirement must be judged in view of this limited scope of the progeny claims. As amended, the claims are drawn to only a limited scope of progeny, progeny which but for Applicant's creation of PH51H could never have existed. This is in harmony with the statement in section 2163 of the MPEP that "the written description requirement promotes the progress of the useful arts by ensuring inventions are adequately described in the specification in exchange for the right to exclude." That quid pro quo of patent law has been met by the Applicant in the present case, and to use written description to deny adequate patent protection would be contrary to the stated purpose of the written description requirement.

The Examiner goes on to reject claims to PH51H plants further containing transgenes and single gene conversions under 35 U.S.C.112, first paragraph. Applicant notes examples of traits and single gene conversions are given in the specification on page 20 lines 16-25 and page 22, line 18 through page 32, line 4. Even if more than one trait is affected by the transgene, the genetics of PH51H will be only minimally affected. The Examiner must consider all limitations of the claimed invention. While the Examiner is focusing on traits, the Applicant points out that she is not claiming so broadly as to claim any maize plant, regardless of source, comprising those traits. Applicant is claiming PH51H, or a limited set of plants derived there from, that retain significant features of PH51H. Applicant has made an enabling deposit of PH51H with the ATCC,

and Applicant is seeking a fair scope of protection as the quid pro quo for the teaching in the specification and the deposit of the material. The insertion of one or a few genes into a genome that is estimated to have over 50,000 to 80,000 genes (Xiaowu, Gai et al., Nucleic Acids Research, 2000, Vol. 28, No. 1, 94-96) is a minor change to PH51H and will not prevent one of skill in the art from identifying the plant as PH51H. In addition, to expedite prosecution, Applicant has amended claims 11 and 30. Claim 11 now reads, "The maize plant, or parts thereof, of claim 2, wherein the plant or parts thereof have been transformed so that its genetic material contains one or more transgenes that confer a qualitative trait." Qualitative traits, as described in an introductory plant breeding book, are traits that "have phenotypes that can be divided into discrete classes...They are controlled by one or a few major genes whose expression is not influenced markedly by the environment" (Fehr, w., Principles of Cultivar Development, vol. 1, 1987, page 26). Claim 30 now reads, "The maize plant, or parts thereof, of claim 2, wherein the plant or parts thereof, further comprise one or more transgenes, and wherein the morphology and physiology of the maize plant comprising the transgene is substantially the same as inbred maize line PH51H." The Examiner has suggested that claims 11 and 30 be amended to list the types of transgenes contemplated in the specification, for example disease or pest resistance genes, provided the prior art teaches those isolated genes. The Applicant believes that an amendment as suggested by the Examiner is limiting the scope to which the Applicant is entitled. Examples of transgenes are given in the specification. The Examiner states that, "Transgenes may also be of any gene, including those that effect more than one trait. The morphological and physiological characteristics of any such plant are not described. For example, a transgene that is a transcription factor can effect more than just one gene, and multiple traits. Such plants would express different morphological and physiological traits from PH51H, which are not described." Applicant points out that the molecular profile of such a plant would be substantially unchanged therefore one would be able to identify such a plant. Applicant amends with traverse in order to expedite allowance.

The Examiner rejects claims 12, 13, 31, and 32. Claims 12 and 31 are drawn to the method of crossing a PH51H plant containing a transgene with another plant. Claims 13 and 32 are to the plant made from the method. The claims have been amended for clarification purposes. Applicant points out that the methods are fully described. Furthermore, one of ordinary skill in the art would know if they were using or one could easily identify if they were using PH51H or PH51H further containing a

transgene to develop a hybrid. All F1 plants would comprise essentially the same genetic markers as the deposited PH51H. It is well known to anyone skilled in the art that a hybrid has a genome with one set of the alleles from each inbred. Therefore the genetic profile exhibited in the deposit would be exhibited in the hybrid. The plant of claim 13 would have the genetic profile of PH51H except at the site of integration of the transgene. The change of one to a few genes out of an estimated 50,000 to 80,000 genes is a minor change and will not prevent one of ordinary skill in the art from identifying the plant as PH51H. One of ordinary skill in the art would also know how to cross PH51H containing a transgene with another plant to produce a hybrid. Thus, the Applicant has described the invention with sufficient specificity to enable others to make and use the invention. In light of the arguments and amendments, the Applicant requests that the Examiner withdraw his rejection to claims 12, 13, 31, and 32.

The Examiner also rejects claims 37-39 under 35 USC § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor, at the time of the application was filed, had possession of the claimed invention. Claims 37-39 are clearly directed to growing out an F(1) hybrid in which PH51H is a parent and searching for PH51H inbred seed. Due to the imperfect process of seed production parent seed can sometimes be contained in the hybrid seed bag. The claims merely claim the method of searching for inbred PH51H seed within a bag of hybrid seed. The method is also clearly described in the specification on page 5, line 21 through line 7 on page 6. The Applicant requests that the Examiner withdraw his rejection to claims 37-39.

Lastly, The Examiner has rejected certain method claims under written description. Applicant points out that the methods are fully described, as is the starting material in the method, PH51H. One of ordinary skill in the art would know how to cross PH51H to develop an F1 hybrid and also how to self plants derived from crosses with PH51H for the purpose of developing an inbred plant. In *Ex parte Parks*, 30 USPQ 2d 1234 (B.P.A.I. 1994), the Board of Appeals stated, "Adequate description under the first paragraph of 35 U.S.C. 112 does not require *literal* support for the claimed invention. Rather, it is sufficient if the originally-filed disclosure would have conveyed to one having ordinary skill in the art that an appellant had possession of the concept of what is claimed." In *J.E.M. Ag. Supply*, the Supreme Court also acknowledged the value of a utility patent in protecting the use of the line in breeding, when it stated that, "...a breeder can use a plant that is protected by PVP certificate to 'develop' a new inbred line

while he cannot use a plant patented under §101 for such a purpose." Id. at 1873. In light of the amendments to the claims and the foregoing arguments the Applicant requests reconsideration of the rejection under the first paragraph of 35 U.S.C. 112.

11. The Examiner rejects claims 18-20 and 47-49 under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. Applicant traverses the rejection.

The Examiner rejects claims 18-20 and 47-49 that directed to PH51H inbred maize plant further comprising one or more single gene conversions. Claims 18 and 47 are directed to PH51H that contains a gene that has been transferred to PH51H through traditional breeding methods. The claims have been amended to expedite prosecution. Claim 18 now reads, "The maize plant, or parts thereof, of claim 2, further comprising one or more genes that confer a qualitative trait and have been transferred into said maize plant through breeding methods that utilize PH51H as a recurrent parent." Claim 47 now reads, "The maize plant, or parts thereof, of claim 2, further comprising one or more genes that have been transferred into said maize plant by utilizing PH51H as a recurrent parent and wherein the maize plant, or parts thereof, are essentially unchanged from inbred maize line PH51H." Claims 19-20 and 48-49 have been amended for clarification purposes. Once again the Applicant would like to point out that one of ordinary skill in the art would be able to detect a PH51H maize plant that contains genes that have been inserted through crossing. The genetics would be substantially the same as PH51H as would the morphological and physiological traits of PH51H. The specification states, "A further embodiment of the invention is a single gene conversion or introgression of the maize plant disclosed herein in which the gene or genes of interest (encoding the desired trait) are introduced through traditional (non-transformation) breeding techniques, such as backcrossing (Hallauer *et al*, 1988)."

The Examiner has cited articles and states that they "teach that it is unpredictable whether the gene or genes responsible for conferring a phenotype in one plant genotypic background may be introgressed into the genetic background of a different plant, to confer a desired phenotype in said different plant." The Examiner states that, "Hunsperger et al. teach that the introgession of a gene in one genetic background on any plant of the same species, as performed by sexual hybridization, is unpredictable in producing a single gene conversion plant with a desired trait (column 3,

lines 26-46). This is not what is taught by Hunsperger et al. Hunsperger et al. teaches that a gene that results in dwarfism of a petunia plant can be incorporated into other genetic backgrounds of the petunia species (See column 2, line 67 to column 3, lines 1-4). Hunsperger et al. merely discusses that the level of the expression of that gene differed in petunia plants of different genetic backgrounds. Hunsperger et al. succeeded in incorporating the gene into petunia plants of different genetic backgrounds. Therefore Hunsperger et al. demonstrate that one of ordinary skill in the art can use traditional breeding methods to obtain maize plants containing genes that confer a qualitative trait. The specification provides ample disclosure of starting materials such as maize inbred PH51H, a discussion of traditional breeding methods that may be used, and examples of transgenes and naturally occurring genes. Please note in Hallauer et al. (1988) on page 472, submitted in the information disclosure statement, it states that, "For single gene traits that are relatively easy to classify, the backcross method is effective and relatively easy to manage."

The Examiner goes on to state that, "Kraft et al. teach that linkage disequilibrium effects and linkage drag prevent the making of plants comprising a single gene conversion, and that such effects are unpredictably genotypic specific and loci-dependent in nature (page 323, column 1, lines 7-15)." Applicant disagrees that the article states such points. Applicant assumes that the Examiner is trying to point out that one gene cannot be introduced into a plant using traditional breeding techniques such as backcrossing without also introducing closely linked genes into the plant. It is well understood in the relevant art that DNA surrounding the gene of the desired trait is introduced into the plant when traditional breeding techniques are utilized to insert a gene into a plant of interest. It is also understood in the art that introducing a gene into a plant variety such as PH51H is an insubstantial change to the variety. The World Seed Organization, on it's web site, writes, "The concept of an essentially derived variety was introduced into the 1991 Act of the UPOV Convention in order to avoid plagiarism through mutation, multiple back-crossing and to fill the gap between Plant Breeder's Rights and patents." As determined by the UPOV Convention, essentially derived varieties may be obtained for example by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, backcrossing, or transformation by genetic engineering. The commercialization of an essentially derived variety needs the authorization of the owner on the rights vested in the initial variety." International Convention for the Protection of New Varieties

of Plants, as amended on March 19, 1991, Chapter V, Article 14, Section 5(c), (emphasis added). A copy of the relevant portion of the UPOV Convention and the World Seed Organization web site is attached as Appendix C.

An example of how one of ordinary skill in the art can transfer a gene conferring a qualitative trait into a variety through backcrossing is demonstrated by the fact that the commercial market now distributes a multitude of products produced in this manner. Such conversion lines are easily developed without undue experimentation. Poehlman et al. (1995) on page 334, submitted in the information disclosure statement, states that, "A backcross-derived inbred line fits into the same hybrid combination as the recurrent parent inbred line and contributes the effect of the additional gene added through the backcross."

The Examiner goes on to state that, "Eshed et al. teach that in plants, epistatic genetic interactions from the various genetic components comprising contributions from different genomes may effect quantitative traits in genetically complex and less than additive fashion (page 1815, column 1, line 1 to page 1816, column 1, line 1). The Applicant would like to first point out on page 1816, column 1, lines 1-5 of the Eshed et al. article it states, "Recent studies that detected epistasis of selected QTL in Drosophila (Long et al. 1995), soybean (Lark et al.1995) and maize (Doebley et al.1995; Cockerham and Zeng 1996) did not show a less-than-additive trend." Emphasis added. Applicant also adds that transferring a qualitative trait does not require undue experimentation. Please note Hallauer et al. (1988) on page 472, submitted in the information disclosure statement, which states, "For single gene traits that are relatively easy to classify, the backcross method is effective and relatively easy to manage." As stated previously claims 18 and 47 have been amended to expedite prosecution. In claim 18, the genes transferred into PH51H are now limited to qualitative traits. Claim 47 is now limited to plants that are essentially unchanged from PH51H. Given the arguments and the amendments the Applicant requests reexamination and reconsideration of the claims.

As noted in the specification, the development of an inbred line is a time consuming and labor intensive activity. On average, between 10,000 to 20,000 lines are created and screened in order to develop any maize inbred line for which the Applicant files a patent application. Once developed, the inbred line is useful for two purposes: (1) to make commercial hybrids, and (2) as a source of breeding material for the development of new inbreds that retain the original inbred's desired characteristics. A breeder desiring to make a line with similar traits to PH51H would be greatly advantaged

by being able to use PH51H as starting material. This is because the linked genes arranged through Applicant's breeding efforts, and fixed in PH51H, can be maintained in the progeny of PH51H by a breeder of ordinary skill in the art. For example, if a breeder of ordinary skill in the art desired a waxy-kernel corn version of PH51H, the breeder could cross PH51H to a waxy-kernel corn variety, select for progeny with the desirable traits of PH51H that also express the waxy kernel trait, and continue selecting for the traits of PH51H combined with waxy kernels. Optionally, the breeder could backcross to PH51H to obtain further genetic contribution from PH51H. The end result is the development of an inbred line with substantially all of the benefit of Applicant's work but with only a fraction of the effort.

In light of the amendments to the claims and the foregoing arguments the Applicant requests reconsideration of the rejection under the first paragraph of 35 U.S.C. 112.

Claim Rejections under 35 U.S.C. § 102 and 103

12. Examiner states that, "Claims 14, 17, 33, 36, 41, 43, 45, and 46 remain rejected under 35 U.S.C. 102(e) as anticipated by or in the alternative, under 35 U.S.C. 103(a) as obvious over Piper (U.S. Patent No. 6,188,001)." Applicant traverses the rejection.

Applicant has cancelled claims 45 and 46. Applicant has amended claims 14, 17, 33, 36, 40, 41, 42, and 43. Claims 36 and 41 are now one cross away from PH51H. Claim 41 clearly states that PH51H must be used to obtain a PH51H-progeny maize plant. Claim 42 has been amended so that it does not allow any further crosses away from PH51H. Thus claim 42 is the selfing of the plant derived by the one cross away from PH51H made in claim 40. Claim 43 has been amended for clarification purposes. These PH51H-progeny plants are limited to one cross away from PH51H and the progeny plants are limited by the use of PH51H in the initial cross. Applicant contends that progeny of PH51H could not be the same as PH1W0 or the progeny of PH1W0 because PH51H is not PH1W0. One would not be able to obtain plants within one cross of PH51H through modification of the maize inbred PH1W0 taught in U.S. Patent No. 6,188,001 because PH51H comprises a unique and nonobvious combination of genetics. As evidenced by past arguments and the declaration of Stephen Smith submitted as Appendix D, the claimed progeny plants of PH51H retain unique and nonobvious combinations of genetics derived from PH51H. Thus, they deserve to be considered new compositions in their own right.

In light of the above, Applicant respectfully requests that the Examiner reconsider and withdraw the rejection to claims 1-49 under 35 U.S.C. 102 (e) and 103(a).

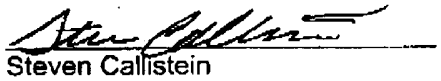
Cancellation of claims 44, 45, 46, 50, and 51 and amendment of claims 3, 4, 11, 12, 13, 14, 17, 18, 19, 20, 22, 23, 30, 31, 32, 33, 36, 40, 41, 42, 43, 47, 48, 49 and 51 does not in any way change the claim scope which the Applicant believes is allowable but is meant to hasten the issuance of the patent.

CONCLUSION

Attached hereto is a marked-up version of the changes made to the specification and claims by current amendment. The attached page is captioned "**VERSION WITH MARKINGS TO SHOW CHANGES MADE**".

Applicant submits that in light of the foregoing amendments and the remarks, the claims 1-43 and 47-49 are in condition for allowance. Reconsideration and early notice of allowability is respectfully requested. If it is felt that it would aid in prosecution, the Examiner is invited to contact the undersigned at the number indicated to discuss any outstanding issues.

Respectfully submitted,
Lori Lisa Carrigan



Steven Callistein
Reg. No. 43,525
Attorney for Applicant

Steven Callistein
Pioneer Hi-Bred International
7100 NW 62nd Avenue
P.O. box 1000
Johnston, IA 50131-1000
(515)-254-2823

09/490,884

VERSION WITH MARKINGS TO SHOW CHANGES MADE

Claims 44, 45, 46, 50, and 51 have been cancelled.

Claims 3, 4, 11, 12, 13, 14, 17, 18, 19, 20, 22, 23, 30, 31, 32, 33, 36, 40, 41, 42, 43, 47, 48, and 49 have been amended as follows.

3. (Twice Amended) The maize plant of claim 2, wherein said plant has been [manipulated to be male sterile] detasseled.

4. (Amended) A tissue culture of regenerable cells or protoplasts from the plant of claim 2.

11. (Amended)The maize plant, or parts thereof, of claim 2, wherein the plant or parts thereof have been transformed so that its genetic material contains one or more transgenes [operably linked to one or more regulatory elements] that confer a qualitative trait.

12. (Amended) A method for producing a first generation (F1) hybrid maize plant that contains in its genetic material one or more transgenes, comprising crossing the maize plant of claim 11 with [either] a second plant [of another maize line, or a non-transformed maize plant of the line PH51H, so that the genetic material of the progeny that result from the cross contains the transgene(s) operably linked to a regulatory element].

13. (Amended) [Maize plants] The first generation (F1) hybrid, or parts thereof, produced by the method of claim 12.

14. (Twice Amended) A maize plant, or parts thereof, wherein at least one ancestor of said maize plant is the maize plant of claim 2, said maize plant

expressing a combination of at least two traits which are not significantly different from PH51H when determined at a 5% significance level and when grown in the same environmental conditions, said traits selected from the group consisting of: a maturity of 94-100 based on the Comparative Relative Maturity Rating System for harvest moisture of grain, female yield, scatter grain resistance, tassel size, pollen shed, hybrid yield, [drydown, heat and seasonal drought tolerance, late season plant health,] stalk lodging resistance, test weight, [grain quality,] plant height, and ear [placement] height [, standability, and adaptability to the Northwest, Northcentral, and Northeast regions of the United States, Northern and Central Europe and Canada].

17. (Amended) [A] The PH51H-progeny maize plant, or parts thereof, produced by the method of claim 15, wherein the method comprises 2 or less crosses to a plant other than PH51H or a plant that has PH51H as a parent.

18. (Amended) The maize plant[s], or parts thereof, of claim 2, further comprising one or more [single gene conversions] genes that confer a qualitative trait and have been transferred into said maize plant through breeding methods that utilize PH51H as a recurrent parent.

19. (Amended) The [single gene conversion(s)] maize plant of claim 18, wherein [the gene] at least one of the genes is a dominant allele.

20. (Amended) The [single gene conversion(s)] maize plant of claim 18, wherein [the gene] at least one of the genes is a recessive allele.

22. (Twice Amended)The maize plant of claim [21] 2, wherein [said plant has been manipulated to be male sterile] genes controlling male sterility have been transferred into said maize plant through crossing, that utilizes PH51H as a recurrent parent, and wherein said plant has essentially the same morphology and physiology of inbred line PH51H other than the trait of male sterility.

23. (Amended) A tissue culture of regenerable cells or protoplasts from the plant of claim 21.

30. (Amended) The maize plant, or parts thereof, of claim [21] 2, wherein the plant or parts thereof [have been transformed so that its genetic material contains one or more transgenes operably linked to one or more regulatory elements] further comprise one or more transgenes, and wherein the morphology and physiology of the maize plant comprising the transgene is substantially the same as inbred maize line PH51H.

31. (Amended) A method for producing a first generation (F1) maize plant [that contains in its genetic material one or more transgenes,] comprising crossing the maize plant of claim 30 with [either] a second plant [of another maize line, or a non-transformed maize plant of the line PH51H, so that the genetic material of the progeny that result from the cross contains the transgene(s) operably linked to a regulatory element].

32. (Amended) [Maize plants] The first generation (F1) maize plant, or parts thereof, produced by the method of claim 31.

33. (Twice Amended) A maize plant, or parts thereof, wherein at least one ancestor of said maize plant is the maize plant of claim [21] 2, [said maize plant expressing a combination of at least two traits which are not significantly different from PH51H when determined at a 5% significance level and when grown in the same environmental conditions, said traits selected from the group consisting of: a maturity of 94-100 based on the Comparative Relative Maturity Rating System for harvest moisture of grain, female yield, scatter grain resistance, tassel size, pollen shed, hybrid yield drydown, heat and seasonal drought tolerance, late season plant health, stalk lodging resistance, test weight, grain quality, plant height, ear placement, standability, and adaptability to the Northwest,

Northcentral, and Northeast regions of the United States, Northern and Central Europe and Canada] and wherein the pedigree of said PH51H-progeny maize plant is within 2 or less crosses to a plant other than PH51H or a plant that has PH51H as a parent.

36. (Amended) [A] The maize plant, or parts thereof, produced by the method of claim 34 wherein the method comprises no more than one cross to a plant other than PH51H or a plant that has PH51H as a parent.

40. (Twice Amended) A method for producing a first generation (F1) PH51H-[derived] progeny maize plant, comprising:

(a) crossing inbred maize line PH51H, representative seed of said line having been deposited under ATCC Accession No. PTA-4261, with a second maize plant to yield progeny maize seed;

(b) growing said progeny maize seed, under plant growth conditions, to yield said first generation (F1) PH51H-[derived] progeny maize plant.

41. (Thrice Amended) A first generation (F1) PH51H-[derived] progeny maize plant, or parts thereof, produced by the method of claim 40.

42. (Twice Amended) The method of claim 40, further comprising:

[(c) crossing said PH51H-derived maize plant with itself or another maize plant to yield additional PH51H-derived progeny maize seed;

(d) growing said progeny maize seed of step (c) under plant growth conditions, to yield additional PH51H-derived maize plants;

(e) repeating the crossing and growing steps of (c) and (d) from 1 to 4 times to generate further PH51H-derived maize plants] selfing said first generation (F1) PH51H-progeny maize plant for

successive filial generations to generate a PH51H inbred progeny maize plant.

43. (Twice Amended) The [further PH51H-derived] PH51H inbred progeny maize plant, or parts thereof, produced by the method of claim 42.

47. (Amended) The maize plant[s], or parts thereof, of claim [21] 2, further comprising one or more [single gene conversions] genes that have been transferred into said maize plant by utilizing PH51H as a recurrent parent and wherein the maize plant, or parts thereof, are essentially unchanged from inbred maize line PH51H.

48. (Amended) The [single gene conversion(s)] maize plant of claim 47, wherein at least one of the [gene] genes is a dominant allele.

49. (Amended) The [single gene conversion(s)] maize plant of claim 47, wherein at least one of the [gene] genes is a recessive allele.

Assessing Probability of Ancestry Using Simple Sequence Repeat Profiles: Applications to Maize Hybrids and Inbreds

Donald A. Berry,*[,1] Jon D. Seltzer,[†] Chongqing Xie,[‡] Deanne L. Wright[‡] and J. Stephen C. Smith[‡]

*The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, [†]Third Wave Technologies, Inc., Madison, Wisconsin 53719 and [‡]Pioneer Hi-Bred International, Inc., Johnston, Iowa 50131

Manuscript received July 24, 2001
Accepted for publication March 11, 2002

ABSTRACT

Determination of parentage is fundamental to the study of biology and to applications such as the identification of pedigrees. Limitations to studies of parentage have stemmed from the use of an insufficient number of hypervariable loci and mismatches of alleles that can be caused by mutation or by laboratory error and that can generate false exclusions. Furthermore, most studies of parentage have been limited to comparisons of small numbers of specific parent-progeny triplets thereby precluding large-scale surveys of candidates where there may be no prior knowledge of parentage. We present an algorithm that can determine probability of parentage in circumstances where there is no prior knowledge of pedigree and that is robust in the face of missing data or mistyped data. We present data from 54 maize hybrids and 586 maize inbreds that were profiled using 195 SSR loci including simulations of additional levels of missing and mistyped data to demonstrate the utility and flexibility of this algorithm.

DETERMINATION of parentage is fundamental to the study of reproductive and behavioral biology. The increasing availability of highly discriminant genetic markers for many diverse species provides the potential to uniquely characterize individuals at numerous loci and to unambiguously resolve parentage where genealogical relationships are unknown, in error, or in dispute.

Identification of parent-progeny relationships in wild populations of animals and plants provides insights into the success of various reproductive strategies (Ellstrand 1984; Smouse and Meagher 1994; Alderson *et al.* 1999) and has allowed for the implementation of management programs to conserve genetic diversity (Miller 1975; Rannala and Mountain 1997). The association of pedigree with physical appearance or performance in domesticated animals and plants allows parents that have contributed favorable alleles for desirable traits through selective breeding programs to be identified (Bowers and Meredith 1997; Sefc *et al.* 1998; Vankan and Faddy 1999). These applications of associative genetics facilitate further progress in genetic improvement through breeding. Establishment of parentage is also useful to secure legal rights of guardianship in humans, to help protect intellectual property in plant varieties, to validate breed pedigrees of domesticated animals, to protect stocks of fish, and to identify provenance of meat that is available in supermarkets (Gotz and Thaller 1998; Primmer *et al.* 2000; White *et al.* 2000).

Most studies of pedigree have utilized exclusion analysis where the molecular marker genotypes of either one or a restricted number of potential triplets of offspring and putative parents are compared. Often the identity of the mother is not in question; the maternal profile is subtracted from that of the offspring and the deduced paternal profile is then compared with candidate father genotypes (Ellstrand 1984; Hamrick and Schnabel 1983). Individuals who could not have contributed the paternal genotype are excluded; the remainder are possible parents. Nonpaternity in humans is generally declared only on the basis of exclusions exhibited by at least two unlinked and independent loci. This criterion of exclusion reduces the likelihood of a false declaration of nonpaternity on the basis of marker results that are actually due to mutation within the phylogeny. Bein *et al.* (1998) show that evidence of nonpaternity should require exclusions at loci on different chromosomes to avoid erroneous conclusions that would be made due to nondisjunction at meiosis leading to uniparental inheritance. A requirement for at least three independent exclusions to declare nonpaternity in humans has also been instituted (Gunn *et al.* 1997). In studies of natural populations of animals or plants where numerous parent-progeny triplets are examined it is usual to accept a single exclusionary event as evidence of nonpaternity (Marshall *et al.* 1998). Paternity testing has been extended to situations where DNA from either parent is unavailable. For example, paternity can still be established in circumstances where the putative father is deceased but his parents are still alive (Helminen *et al.*

[1]*Corresponding author:* Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Box 447, Houston, TX 77030-4009. E-mail: dberry@mdanderson.org

Genetics 161: 813–824 (June 2002)

1991; Bockel *et al.* 1992). Chakraborty *et al.* (1994) demonstrate that paternity can be determined in cases where the mother is unavailable for testing. Lang *et al.* (1993) partially reconstructed the DNA profile of a missing crocodile parent using profiles of the mother and progeny.

Chakraborty *et al.* (1988) and Smouse and Meagher (1994) report that reliance upon exclusion alone has usually failed to unambiguously resolve paternity. Limitations have stemmed from the use of an insufficient number of independent hypervariable loci. Other statistical methods are therefore required to calculate the likelihood of paternity for each nonexcluded male (Berry and Geisser 1986; Meagher 1986; Meagher and Thompson 1986; Thompson and Meagher 1987; Devlin *et al.* 1988; Berry 1991). Marshall *et al.* (1998) draw attention to the quality of data that is encountered practically in genotypic surveys. Maternal genetic data may or may not be available, data may be absent for some candidate males, data may be missing for some loci in some individuals, null alleles exist, and typing errors occur. Reconstructing or validating the pedigrees of varieties of cultivated plants often provides additional challenges because their phylogenies can reveal apparent exclusions that masquerade as non-Mendelian inheritance. For example, apparent exclusions can occur in circumstances where an individual is used as a parent prior to completion of the inbreeding process. The development of parent and progeny then continue on parallel but separate tracks thereby allowing the possibility that alleles that are subsequently lost through inbreeding in the parent can still become fixed in the progeny. It is also possible to create many offspring from a single mating and to use the same parent repeatedly in "backcrossing." Therefore, many individual inbred lines, varieties, or hybrids can be highly related. In consequence, there are numerous (and often very similar) pedigrees. The effective number of marker loci that can discriminate between alternate pedigrees is proportionally reduced as parents are increasingly related. Consequently, inbred lines can be more similar to one or more sister or other inbreds than those inbreds are to one or both of their parents.

It has not been usual to search among hundreds of individuals to identify the most probable maternal and paternal candidates for a specific progeny. Most studies of parentage are in circumstances where there is *a priori* information for at least one of the parents (usually the maternal parent). Limited availability of marker loci and the lack of very high-throughput genotyping systems offering inexpensive datapoint costs may have focused research on studies that involve relatively few individuals and where there is at least some *a priori* indication of parentage. Studies that have been conducted without *a priori* information on parentage include species where reproductive behavior renders identification of the maternal parent difficult or impossible. Examples include those undertaken on birds that practice brood parasitism (Anderson *et al.* 1999) or extra-pair copulation (Wetton *et al.* 1992) or on species such as the wombat that are difficult to observe in the wild (Taylor *et al.* 1997).

Two circumstances favor a revised approach to the statistical analysis of pedigree. First, molecular marker technologies are rapidly developing and will allow numerous loci to be typed for thousands of individuals rapidly and inexpensively. A greater number and diversity of larger-scale studies of pedigree can be expected within the plant and animal kingdoms including individuals in which there is no prior knowledge of pedigree. A larger number of markers mean a greater chance for errors. Therefore, the second circumstance follows: Procedures that are efficient and robust in the face of apparent exclusions, missing data, and laboratory error are required.

The purpose of this article is to describe and evaluate a methodology that can be used to quantify the probability of parentage of hybrid genotypes. We focus on parentage because it is the primary focus of published literature and it is the easiest level of ancestry to understand. The method is robust in the face of mutation, pseudo-non-Mendelian inheritance (apparent exclusions) due to residual heterozygosity in parental seed sources, missing data, and laboratory error. The methodology has a number of advantages: (i) It can accommodate large datasets of possible ancestors (hundreds of inbreds or hybrids each profiled by >100 marker loci), (ii) it does not require prior knowledge about either parent of the hybrid of interest, (iii) it does not require independence of the markers, and (iv) it can successfully discriminate between many highly related and genetically similar genotypes. We demonstrate the effectiveness of this approach to identify inbred parents of maize (*Zea mays* L.) hybrids using simple sequence repeat (SSR) marker profiles for 54 maize hybrids together with their parental and grandparental genotypes included among a total of 586 inbred lines. The methodology is applicable to the investigation of parentage for all progeny developed from parental mating without subsequent generations of inbreeding.

MATERIALS AND METHODS

Algorithm: Consider an index hybrid whose parentage is unknown or in dispute. Inbreds in an available database are possible ancestors of the hybrid. The objective is to find the probabilities of closest ancestry for each inbred on the basis of information from SSRs from the index hybrid and the inbreds. There is no reason to trim the database by removing inbreds thought to be unrelated to the index hybrid because their lack of relationship will be discovered.

Consider a pair of possible ancestors, inbred i and inbred j. There is nothing special about this particular pair as all pairs will be treated similarly. The process involves calculating the probability that inbreds i and j are in the hybrid's ancestry, repeating this for all pairs of inbreds in the database.

The basis of the algorithm is Bayes' rule (*e.g.*, BERRY 1991, 1996). Let $P(i, j|\text{SSRs})$ stand for the (posterior) probability that i and j are ancestors of the index hybrid given the information from the various SSRs. Let $P(i, j)$ stand for the unconditional (or prior) probability of the same event. Finally, $P(\text{SSRs}|i, j)$ is the probability of observing the various SSR results if in fact i and j are ancestors. Bayes' rule says

$$P(i, j|\text{SSRs}) = P(\text{SSRs}|i, j) \times P(i, j)/\sum[P(\text{SSRs}|u, v) \times P(u, v)],$$

where the sum in the denominator is over all pairs of inbreds, indexed by u and v. $P(\text{SSRs}|i, j) \times P(i, j)$ is one of the terms in the denominator. (To compute the denominator in the above expression, fix a particular order to the inbreds in the database and take $u < v$ in expressions involving the pair (u, v). If there are 586 inbreds, for example, then the number of pairs and the number of terms in the denominator is $586(587)/2 = 171{,}991$.) Inbreds i and j may be parents or grandparents or other types of relations or bear no relationship at all to the hybrid. If there are more than two ancestors in the database, such as both parents and all four grandparents, then the possible pairs involving these ancestors will generally have the highest posterior probabilities. If the hybrid's true parents are in the database, then as a pair they will typically have the highest overall posterior probability. If both i and j happen to be related to one particular parent of the hybrid, then as a pair their posterior probability will be low because they will not usually account for many of the alleles that are contributed by the other parent of the hybrid.

We will make the "no-prior-information" assumption that $P(u, v)$ is the same for all pairs (u, v). This implies that this factor is cancelled from both numerator and denominator in the above expression, giving:

$$P(i, j|\text{SSRs}) = P(\text{SSRs}|i, j)/\sum P(\text{SSRs}|u, v).$$

The problem is then to calculate a typical $P(\text{SSRs}|i, j)$. Assume inbreds i and j are both ancestors. We calculate the probability of observing the resulting hybrid under this assumption. We make no assumptions about relationships among the various inbreds. Other possible ancestors will be considered implicitly in the calculation by allowing their alleles to be introduced through breedings with i and j. However, the nature of such breedings is not specified. Suppose inbred i's alleles are (a, b). Each descendant of inbred i receives one of these two alleles or not. An immediate descendant receives one with probability 1 (barring mutations). A second generation descendant receives one of them with probability 0.5. And so on. Since degree of ancestry (if any) is unknown, we label the actual probability of passing on one of these alleles to be P. Similarly, an allele from inbred j has been passed down to the hybrid or not, and the probability of the former is P. In the following, P will be taken to equal 0.50, although we will also consider $P = 0.99$ in some of the calculations.

Assuming $P = 0.50$ is consistent with the closest ancestors in the database being grandparents. However, we are not interested in grandparents *per se*. If the closest ancestors in the database were parents, then as indicated above P should equal 1 (ignoring mutations and laboratory errors). Our primary concern is when the parents are not in the database. In this case P is no greater than 0.50. Assuming $P = 0.50$ is robust over the middle range of possible values of P. One way in which it is robust is if there may be mutations and laboratory errors, in which case P would have to be < 1. Taking P to equal 0.50 levies little penalty against a particular pair in which there is an apparent exclusion from direct parentage. Therefore taking P to be < 1 means that if the true parents are in the database then they will not be ruled out if there happen to be mutations and laboratory errors. And if the closest ancestors in the database are more remote than grandparents, they are likely to be identified because they will usually have the fewest mismatches of the lines considered.

When i and j are ancestors there are four possibilities: (1) The alleles of both inbreds i and j were passed to the hybrid. (2) inbred i came through but not inbred j. (3) inbred j came through but not inbred i, and (4) neither inbred came through. Assuming independence, these have respective probabilities P^2, $P(1 - P)$, $P(1 - P)$, $(1 - P)^2$. In the case $P = 0.50$, all of these probabilities equal 0.25.

An instance of the law of total probability (Sec. 5.3, BERRY 1996) is that the probability of observing a hybrid's alleles is the average of the conditional probability of this event given the above four cases. The simplest of the four cases is the first possibility: Assuming the hybrid's alleles are passed down directly from both inbreds, the probability of observing the hybrid's genotype is either 1 or 0 depending on whether the hybrid shares both inbreds' alleles. (It is especially easy when both inbreds are homozygous.) The other three cases require an assumption regarding the possibility that an inbred's allele is not passed to the hybrid but is interrupted by a mutation, a laboratory error, or intervening breeding. We regard such an allele as being selected from all known alleles with probability 1/(number of alleles), where the number of alleles is the total number of alleles known to exist at the locus in question. An alternative approach would be to use the allelic proportions that are present in the database (or in another database). However, the lines in the database may not be randomly selected from any population. For example, a line that has been highly used in breeding would have many derivative lines in the database, in which case the frequencies of its alleles will be artificially inflated. Assuming equal probabilities for the various alleles at a given locus is robust in the sense that it is not affected by adding and dropping lines from the database.

There are many cases to consider when computing the probability of observing a hybrid's alleles, depending on the zygosity of the hybrid and the inbreds, and allowing for the possibility of missing alleles or "extra alleles" in the assessment of the hybrid and inbred genotypes. These possibilities are too numerous to list. Instead we give three simple examples. All the examples have homozygous inbreds, the most common case. And each of the three hybrids has two alleles, again the most common case. We suppose that the measured alleles for three SSRs and a particular trio of hybrid and ancestor inbreds are as we have indicated in Table 1.

For SSR 1 there are three known alleles, one in addition to alleles *a* and *b* that are listed for the three lines (hybrid, inbred i, and inbred j) in Table 1. For SSR 2 and SSR 3 there are two known alleles in addition to those listed. The calculations in the right half of Table 1 will now be explained. Implicit in calculating $P(\text{SSR}|i, j)$ is the assumption—required in both the numerator and denominator of Bayes' rule—that inbreds i and j are ancestors of the hybrid. Consider SSR 1. In case 1 above, both ancestors' alleles (as measured by the laboratory process) are assumed to pass to the index hybrid, and so in this case the hybrid is necessarily *ab*. The probability of observing the actual hybrid's genotype is 1 for case 1, as shown in Table 1. In case 2, we assume that inbred i's allele passes to the hybrid but inbred j's does not. Indeed, the hybrid has an *a* allele. The probability of observing a *b* as the other allele is 1/(number of alleles) = 1/3, as shown in Table 1. Case 3 is similar. In case 4, neither ancestor allele is passed to the hybrid; the probability of observing the hybrid's genotype (or any heterozygous genotype) is $2(1/3)(1/3) = 2/9$. Since $P = 0.50$, the overall (unconditional) probability in the rightmost column (17/36) is the simple average of the four cases, as indicated in Table 1.

For SSR 2 and SSR 3 the calculations are similar. For SSR 2 there is some evidence against pair (i, j) being ancestors.

TABLE 1

Probability of observing a hybrid's alleles using three sample SSRs and four possible combinations (cases) of alleles passed, assuming that inbreds i and j are ancestors of the hybrid

SSR	No. of alleles	Hybrid	Inbred i	Inbred j	Probability of observing the hybrid's genotype: Case 1 i, j	Case 2 i, not j	Case 3 not i, j	Case 4 not i, not j	Overall probability $P(\text{SSR}\|i, j)$
1	3	*ab*	*aa*	*Bb*	1	1/3	1/3	2/9	17/36
2	5	*bd*	*bb*	*Cc*	0	1/5	0	2/25	7/100
3	6	*ab*	*cc*	*Dd*	0	0	0	2/36	2/144

SSR, simple sequence repeat marker profile.

but it is not conclusive. For SSR 3 there is even less evidence favoring pair (i, j). It would not take many SSRs with evidence similar to that for SSR 3 to essentially rule out this pair—provided that other pairs are not similarly inconsistent.

To find the overall $P(\text{SSRs}|i, j)$, multiply the individual $P(\text{SSR}|i, j)$ over the various SSRs. There are purely computational issues to address. Each $P(\text{SSR}|i, j)$ is a number between 0 and 1. When there are a great many SSRs, the product of these numbers will be vanishingly small. To lessen problems with computational underflow, for each SSR we multiply $P(\text{SSR}|u, v)$ by the same constant for each pair (u, v); the inverse of the largest possible such probability. For example, since 17/36 is the largest probability for a heterozygous hybrid at an SSR having three alleles (as is the case for SSR 1 in Table 1), we multiply all factors $P(\text{SSR}|u, v)$ by 36/17. To eliminate remaining problems with underflow, we do calculations using logarithms (adding instead of multiplying) and take antilogs at the end.

The probability $P(\text{SSR}|u, v)$ is calculated for all (u, v) pairs and summed over all possible pairings in the database, including that for the inbred pair under consideration: (i, j). This gives the denominator in the expression for $P(i, j|\text{SSRs})$.

To determine the probability that any particular inbred, say inbred i, is the closest ancestor of the index hybrid, sum $P(\text{SSR}|i, v)$ over all inbreds v with $v \neq i$. Call this $P(i|\text{SSRs})$. The maximum of $P(i|\text{SSRs})$ for any inbred i is 1. But since there is one closest ancestor on each side of the family, the sum of $P(i|\text{SSRs})$ over all inbreds i is 2. If there is a particular pair (i, j) for which $P(i, j|\text{SSRs})$ is close to 1 then both $P(i|\text{SSRs})$ and $P(j|\text{SSRs})$ separately will be close to 1.

SSR data: DNA was extracted from 54 maize hybrids and from 586 maize inbreds. All of the hybrids and most inbreds are proprietary products of Pioneer Hi-Bred International; some important publicly bred inbred lines were also included. The inbred parents and grandparents of each hybrid were included within the set of inbreds. Other inbreds that were genotyped include many that are highly related by pedigree to parents and grandparents of the hybrids. The hybrids were chosen because each has a pedigree that is known to us and collectively they represent a broad array of diversity of maize germplasm that is currently grown in the United States ranging from early to late maturity.

A total of 195 SSR loci were used in this study following procedures described in Smith *et al.* (1997), but modified as described below. SSR loci were chosen on the basis that they individually have been shown to have a high power of discrimination among maize inbred lines and collectively they provide for a sampling of diversity for each chromosome arm. Of these SSR loci, the following numbers (in parentheses) were located on individual maize chromosomes as follows: 1 (35), 2 (26), 3 (22), 4 (20), 5 (16), 6 (9), 7 (6), 8 (18), 9 (12), and 10 (14); 17 SSR loci have not yet been mapped. The correlations among the loci are unknown and are irrelevant for our methodology.

Sequence data for primers that allow many of these (and other) SSR loci to be assayed are available at website http://www.agron.missouri.edu. All primers were designed to anneal and amplify under a single set of conditions for PCR in 10-μl reactions. Genomic DNA (10 ng) was amplified in 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-Cl (pH 8.3) using 0.3 units AmpliTaq Gold DNA polymerase (PE Corporation) oligonucleotide primer pairs (one primer of each pair was fluorescently labeled) at 0.17 μM and 0.2 mM dNTPs. This mixture was incubated at 95° for 10 min (hot start); amplified using 45 cycles of denaturation at 95° for 50 sec, annealing at 60° for 50 sec, extension at 72° for 85 sec; and then terminated at 72° for 10 min. A water bath thermocycler manufactured at Pioneer Hi-Bred International was used for PCR reactions. PCR products were prepared for electrophoresis by diluting 3 μl of each product to a total of 27 μl using a combination of PCR products generated from other loci for that same maize genotype (multiplexing) and/or dH20. Dilution of 1.5 μl of this mixture to 5 μl with gel loading dye was performed; it was then electrophoresed at 1700 V for 1.5 hr on an ABI model 377 automated DNA sequencer equipped with GENESCAN software v. 3.0 (PE-Applied Biosystems, Foster City, CA).

PCR products were sized automatically using the "local Southern" sizing algorithm (Elder and Southern 1987). After sizing of PCR products using GeneScan, alleles were assigned using Genotyper software (PE-Applied Biosystems). Generally, allele assignations for each locus were made on the basis of histogram plots consisting of 0.5-bp bins. Breaks between the histogram plots of >1 bp were generally considered to constitute separation between allele bins; however, other criteria, such as the presence of the nontemplate-directed addition of adenine (+A addition) and naturally occurring 1-bp alleles, were used on a marker-by-marker basis to define the allele dictionary. All allele scores were made without knowing the identities of the maize genotypes.

RESULTS

Table 2 presents the probability of closest ancestry of the top five ranking inbred lines for each of 5 hybrids at $P = 0.50$ (Table 2A) and $P = 0.99$ (Table 2B). Probabilities of ancestry are shown for all 54 hybrids and the top ranking inbreds in Figure 1: $P = 0.50$ (Figure 1a) and $P = 0.99$ (Figure 1b). Results for the hybrids presented in Table 2 are featured at the top of Figure 1.

TABLE 2

Probability of ancestry of five hybrids using data obtained from 50, 100, and 195 SSR loci

	50 loci			100 loci			195 loci		
Hybd.	Inbd.	Prob.	SE	Inbd.	Prob.	SE	Inbd.	Prob.	SE
				A. Assuming $P = 0.50$					
3417	SP1	0.9607	0.0125	P1	0.8749	0.0252	P1	1.0000	E-07
	P2	0.8077	0.1963	P2	0.8141	0.2235	P2	0.9957	0.0033
	D2P2	0.1016	0.1038	D1P2	0.1859	0.2235	D1P2	0.0043	0.0033
	D1P2	0.0907	0.0927	SP1	0.1243	0.025	D2P2	E-06	E-06
	P1	0.032	0.0125	D1P1	0.0009	0.0002	SP1	E-06	E-07
3525	P1	0.8545	E-07	P1	0.9999	<E-20	P1	1.0000	<E-20
	P2	0.8188	E-07	P2	0.5437	<E-20	P2	0.9635	0.0528
	D1P2	0.1699	E-07	D1P2	0.4563	<E-20	D1P2	0.0365	0.0528
	GP1	0.1441	E-07	GP1	E-07	E-18	SP1	E-15	<E-20
	GP2	0.0110	E-08	SP1	E-07	<E-20	GP2	E-16	<E-20
3556	P1	1.0000	E-06	P1	0.9999	E-10	P1	1.0000	<E-20
	P2	0.9616	E-08	P2	0.9997	E-10	P2	1.0000	<E-20
	D1P2	0.0340	E-10	D1P2	0.0003	E-14	D1P2	E-09	<E-20
	GP2	0.0043	E-09	D2P2	E-05	E-15	D2P2	E-14	<E-20
	D2P2	0.0002	E-10	D3P2	E-06	E-17	GGP2	E-17	E-17
3905	D1P1	0.9822	E-08	D1P1	0.9803	0.0038	P1	1.0000	E-08
	SP2	0.4927	E-07	SP2	0.6280	0.0976	D1P2	1.0000	E-06
	D2P2	0.2836	E-07	D1P2	0.2321	0.0617	D2P2	E-06	E-06
	D1P2	0.1622	E-07	D2P2	0.1317	0.0372	P2	E-07	E-13
	P2	0.0565	E-07	P1	0.0197	0.0058	D3P2	E-10	E-16
3940	P2	0.9997	0.0001	P2	0.9999	E-05	P2	1.0000	E-09
	D1P2	0.9203	0.0009	P1	0.9970	0.0011	P1	1.0000	E-09
	P1	0.0648	E-05	D1P2	0.0030	0.0011	D1P2	E-11	E-11
	D1P1	0.0127	E-05	D2P2	0.0001	E-05	DP1P2	E-17	E-17
	DP1P2	0.0014	0.0009	DP1P2	0.0001	E-07	D2P2	E-19	E-18
				B. Assuming $P = 0.99$					
3417	SP1	0.9995	0.0001	P1	0.9999	E-05	P1	0.9999	E-08
	P2	0.8836	0.1658	P2	0.9938	0.0107	P2	0.9999	E-08
	D1P2	0.0722	0.1029	D1P2	0.0061	0.0107	D1P2	E-11	E-11
	D2P2	0.0441	0.0628	D1P1	E-05	E-06	D2P2	E-14	E-14
	P1	0.0004	0.0001	SP1	E-05	0	SP1	E-20	E-21
3525	P1	0.9999	0	P1	0.9999	0	P1	1.0000	0
	P2	0.8991	0	D1P2	0.9749	0	P2	0.6135	0.4446
	D1P2	0.1008	E-11	P2	0.025	0	D1P2	0.3864	0.4446
	GP1	E-05	0	D2P2	E-20	0	GP2	E-48	0
	GP2	E-06	E-17	SP1	E-24	0	D2P2	E-49	0
3556	P1	1.0000	0	P1	1.0000	0	P1	0.9999	0
	P2	0.9996	0	P2	0.9999	0	P2	0.9999	0
	D1P2	0.0003	0	D1P2	E-09	0	D1P2	E-22	0
	D1P1	E-11	0	D3P1	E-21	0	D2P1	E-49	0
	D2P1	E-13	0	D2P1	E-21	0	D3P1	E-54	0
3905	D1P1	0.9999	0	D1P1	0.9999	E-08	P1	1.0000	E-09
	P2	0.9992	0	P2	0.9999	E-06	P2	0.9947	E-09
	SP2	0.0006	0	D1P2	E-06	E-06	D1P2	0.0052	E-11
	D1P2	E-05	0	SP2	E-07	E-13	D2P2	E-18	E-18
	D2P2	E-06	0	D2P2	E-09	E-10	D1P1	E-25	E-25
3940	P2	0.9999	E-08	P2	1.0000	E-08	P1	1.0000	E-09
	D1P2	0.9999	E-08	P1	0.9999	E-05	P2	1.0000	E-09
	P1	E-06	E-13	D1P2	E-05	E-05	D1P2	E-24	E-24
	D1P1	E-08	E-13	D2P2	E-12	E-11	DP1P2	E-44	E-44
	DP1P2	E-12	E-12	DP1P2	E-21	E-21	D2P2	E-50	E-49

Hybd., hybrid; Inbd., inbred; Prob., probability; SE, standard error, referring to the variability in the results of the runs; P1, parent one; P2, parent two; SP1, SP2, full sibling of parent one/parent two; D*P1/D*P2, derivatives of parent one/parent two, index * for distinct inbred lines; DP1P2, derivatives of both parent one and parent two.

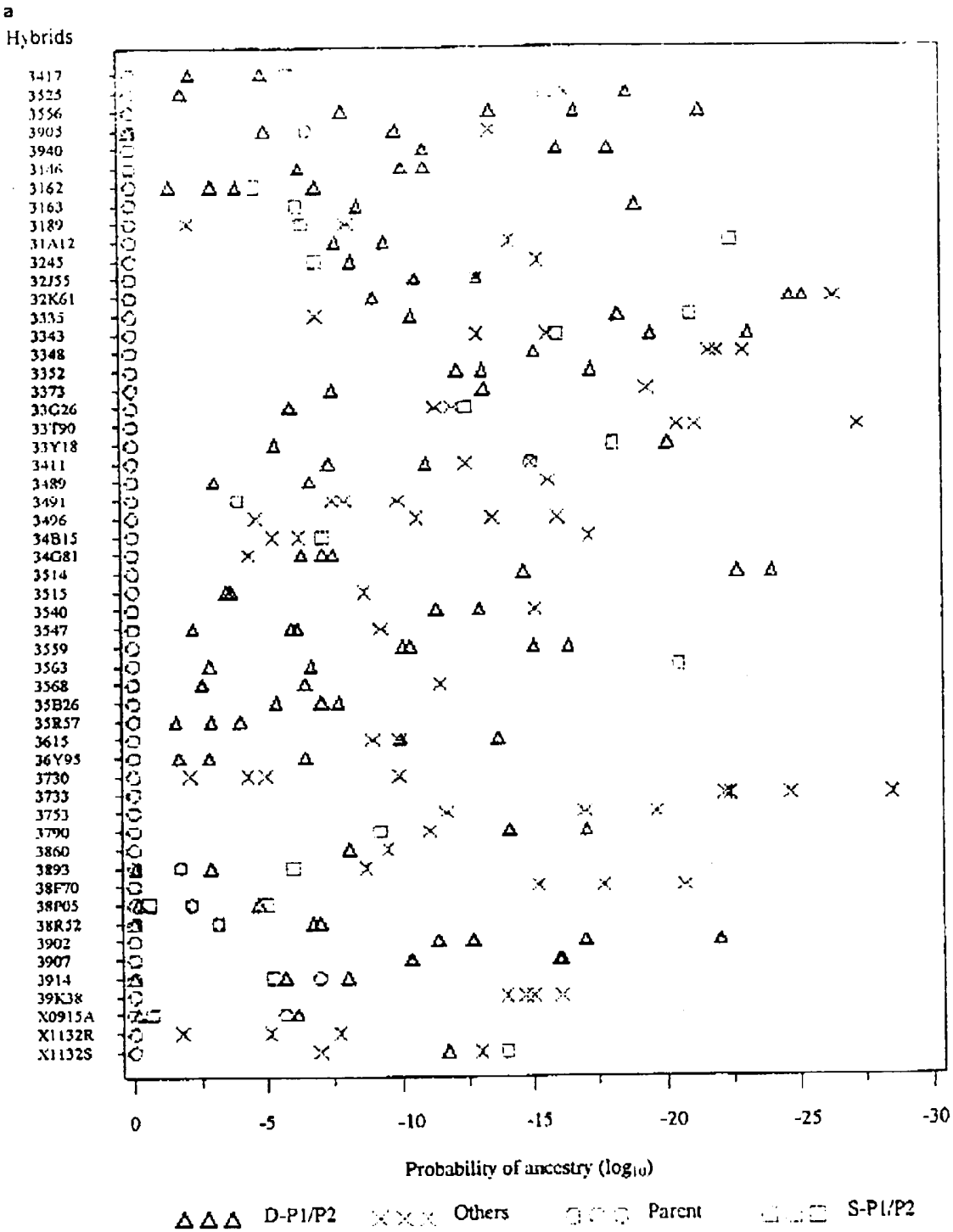


FIGURE 1.—(a) Probabilities of ancestry, assuming $P = 0.50$, for all 54 hybrids and top ranking inbreds—those with probability of ancestry at least 10 %. (b) Probabilities of ancestry, assuming $P = 0.99$, for all 54 hybrids and top ranking inbreds—those with probability of ancestry at least 10 %.

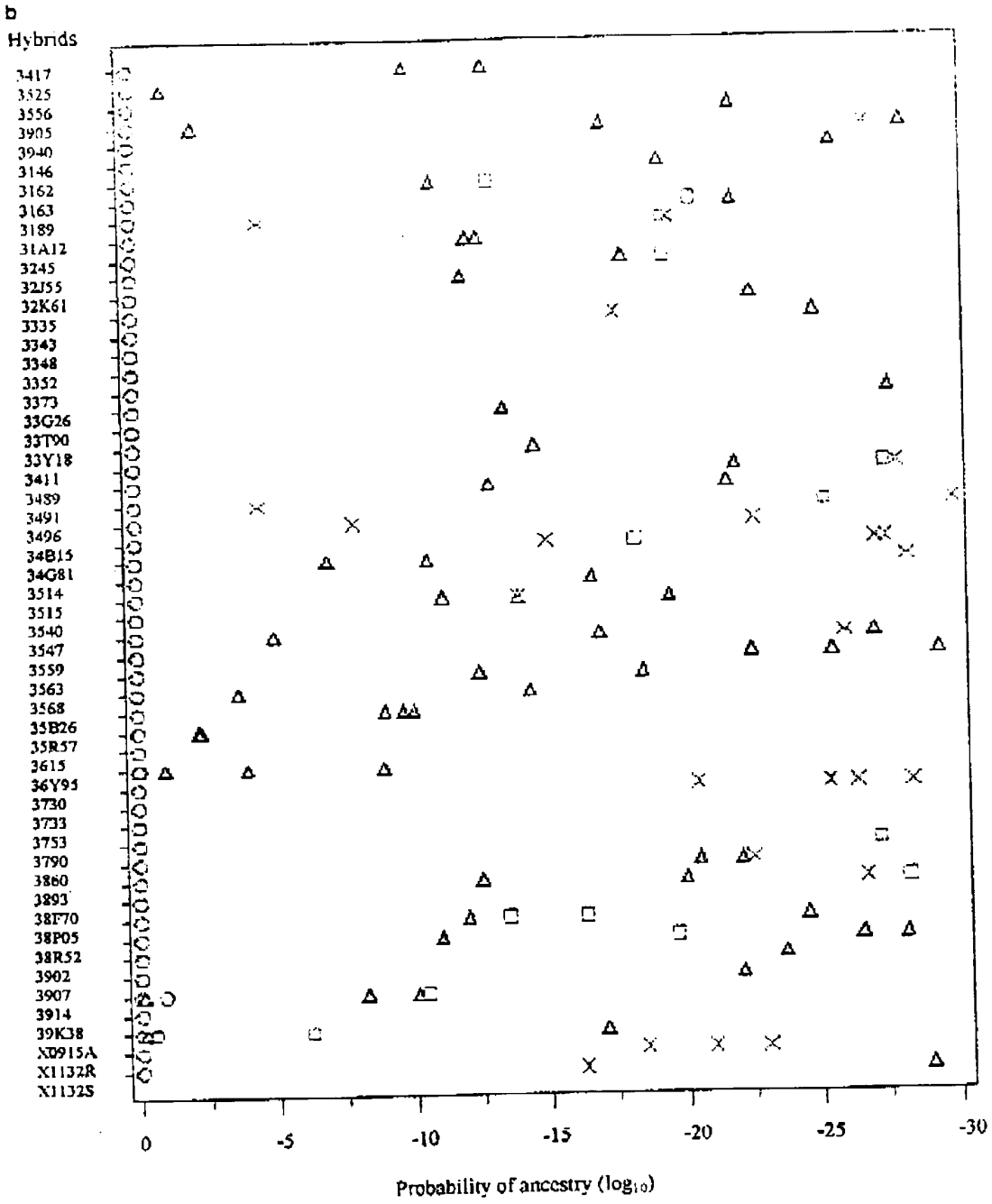


FIGURE 1.—*Continued.*

When the algorithm used $P = 0.50$, the two correct parents were identified as highest in probability for 48 (89%) hybrids (Figure 1). For each of 6 hybrids (3893, 38P05, 38R52, 3905, 3914, and X0915A), one parent ranked in the top two places. The other parent was supplanted either by a sister inbred or by an inbred that was a direct progeny of that parent. Overall, 102 (94%) of 108 parental inbreds were correctly identified. For hybrids where both parents ranked first or second, the range of probabilities for parental lines that ranked first from among all other inbreds ranged from 1.0000 to 0.9997; parental lines ranking second ranged from

1.0000 to 0.9653. For 35 hybrids, both parents had probabilities of ancestry in excess of 0.999. Probabilities of ancestry for nonparents that ranked in first or second places were from 0.9999 to 0.7054. For the majority of hybrids, the probability of the third and highest ranked nonparental inbred was at or below E-06. This indicates that there is usually very little uncertainty about closest ancestors.

When the algorithm used $P = 0.99$ to examine each of the 54 hybrids, both parents were correctly identified for 52 (96%) of hybrids and for 98% (102/104) of the parents across all hybrids (Figure 1). Two hybrids (3914 and X0915A), in which one parent was not ranked in the top two, were also in the subset not ranked in the top two assuming $P = 0.50$ (above). In both cases their ranks improved (both to third rank) and the actual parent was supplanted by an inbred that was a direct progeny of the corresponding parental line. For 49 hybrids, both parents had probabilities of ancestry in excess of 0.999. Among the 5 hybrids having a parent ranking second with a probability of ancestry below 0.999, the lowest of these probabilities was 0.8976 and the highest probability for a third ranking nonparent was 0.1023. For most hybrids the probability for the third and highest ranked nonparental inbred was at or below E-10.

Table 2 also addresses data analysis in circumstances where heterozygous loci occur in inbred lines or where a hybrid is scored for the presence of more than two alleles per locus. The presence of more than a single allele per locus in inbred lines is an infrequent occurrence in well-maintained inbred development and seed increase programs but is possible because ~3–5% of loci can still be segregating and unintended pollination from genotypes not designated as parents of the hybrid can occur. For hybrids, more than two alleles per locus can be scored when DNA is extracted from a bulk of individual plants and because inbred parents are not homozygous due either to residual heterozygosity or to contamination or because one or more direct parents of the hybrid are themselves hybrids. The presence of more than one allele per locus in an inbred line and more than two alleles per locus in a hybrid therefore can be accommodated by multiple runs of the algorithm, each with a random choice of two alleles per locus. Consequently, standard errors in the case of analyzing data from 195 loci tend to be very small because there were few loci where an inbred or hybrid sample (from a bulk of individual plants) was scored for more than two alleles.

Marshall *et al.* (1998) have drawn attention to errors that can be encountered in genotyping surveys. These errors include missing data, null alleles, and typing errors. We therefore investigated the robustness of the algorithm by examining the effects of modifications in the data for five hybrids (3417, 3525, 3556, 3905, and 3940). First, we reduced the number of SSRs used, from the full set of 195 to 100 and then to 50 (Table 2). Use of 50 loci generated incorrect rankings of one parent for each of two hybrids (3417 and 3940) and for both parents of one hybrid (3905). All of these most highly ranked nonparental inbreds were closely related to the true parents for each of the respective hybrids; six different inbred lines were involved. Four were direct progeny of the true parents (one with additional backcrosses from the true parent) and two were full sisters (from a cross of highly related inbreds) of the actual parent of the hybrid. Using 100 loci resulted in correct parental rankings for all hybrids except for 3905 where neither parent ranked in first or second place. Four inbreds outranked the true parents of 3905. All four nonparents were closely related to the respective true parents; three were direct progeny of the true parent of the hybrid (one with additional backcrossing to that parent) and one was a full sister of the true parent. Use of data from all 195 loci corrected the placement for one of the parents of hybrid 3905. Two inbreds that were not parents of this hybrid remained ranked more highly than one of the true parents. Both were direct progeny of that parent, and one of these inbreds had additional backcrossing to that parent in its pedigree.

To address the consequences of laboratory and other sources of error, we artificially compromised data quality beyond the level originally provided by eliminating specific proportions of alleles that had been scored (establishing scenarios where various numbers of SSR alleles were not scored) and by misscoring other alleles (establishing scenarios where various numbers of SSR alleles were scored incorrectly). We also combined the scenarios of missing data and wrongly scored data. Table 3 contains a summary of the results of making these modifications in the data. For all modifications we used data from all SSR loci and we also randomly chose SSR loci to create subsets of 50 and 100 loci. In each case, the program was run 20 times for each hybrid/set of loci. When all 195 loci were examined, replications differed only according to the particular choice of alleles for loci where more than two alleles had been scored.

To evaluate robustness in the face of missing data or mistyped data, we simulated individual and combined categories of these data in the hybrid and all inbred lines at levels of 2, 5, 10, and 25% of the alleles for each of five hybrids and all inbreds beyond the level of error as originally scored by the laboratory. We examined the effects of these levels and types of error for three sizes of database: 50 loci, 100 loci, and all 195 scored loci. The same five hybrids considered in Table 2 were investigated: 3417, 3525, 3556, 3905, and 3940. One of these hybrids (3905) was chosen because one of its parents did not rank among the top two places even when the complete and unmodified data from all SSR loci were used.

Examples of robustness in the face of additional error

TABLE 3

Number of parents ranked in first and second positions (maximum is 2)

Type of simulated data	% level change	No. of loci:	Hybrid 3417 50	3417 100	3417 195	3525 50	3525 100	3525 195	3556 50	3556 100	3556 195	3905 50	3905 100	3905 195	3940 50	3940 100	3940 195	Mean % max.
Missing	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	1	2	2	2	2	0	0	1	1	2	2	73
	5		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	10		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	25		0	2	2	1	1	1	2	2	2	0	0	1	1	1	1	57
	Mean % max.		40	100	100	90	80	90	100	100	100	0	0	50	50	90	90	
Miscored	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	5		1	2	2	2	1	2	2	2	2	0	0	1	1	2	2	73
	10		1	1	2	2	2	2	2	2	2	0	0	1	1	2	2	73
	25		1	0	2	1	1	2	1	2	2	0	1	1	1	2	2	63
	Mean % max.		50	70	100	90	80	100	90	100	100	0	10	50	50	100	100	
Missing plus miscored	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	5		1	2	2	2	1	1	2	2	2	0	0	1	1	2	2	70
	10		1	2	2	1	2	2	2	2	2	0	1	1	1	2	2	77
	25		0	1	2	0	0	0	0	0	0	0	0	1	0	0	1	17
	Mean % max.		40	90	100	70	70	70	80	80	80	0	10	50	40	80	90	
Overall mean			48	87	100	83	77	87	90	93	93	0	7	50	47	90	93	

Hybrids considered are the same as those in Table 2.

for five hybrids using subsets of 50 and 100 loci and all loci are shown in Table 3 where numbers of parents ranking into the top two places are presented. Degradation in the preferential ranking of parent inbreds at a level of 25% additional missing data was shown for one hybrid (3525) with usage of 50, 100, or all SSR loci. Degradation in the preferential ranking of parent inbreds at a level of 25% additional misscored data was shown for hybrid 3556. When both additional levels of missing and misscored data were simulated, degradation in the ability to preferentially rank inbred parents occurred for all hybrids and for all sets of SSR (50, 100, and 195 loci) except for hybrid 3417 when data from 195 SSR loci were used. Over all five hybrids, use of 100 loci improved robustness from the use of 50 loci; use of 195 loci further improved robustness for four hybrids (3417, 3525, 3905, and 3940). The degree of improvement was small, except for hybrid 3905.

We also ranked inbreds according to their probability of ancestry of hybrids when both parents and all inbred derivatives and full-sister inbreds of the respective inbred parents for each hybrid were excluded from the analysis. The results are too voluminous to present here but can be summarized as follows: Using $P = 0.50$, a grandparent of each respective hybrid ranked into first place for 41 (76%) hybrids; probabilities ranged from 0.4976 to 1.0 and most were above 0.9999. Other classes of inbreds that ranked in first position for probability of ancestry were inbreds derived directly by pedigree from a grandparent of the respective hybrid (DGP) for 13% of hybrids, inbreds derived directly by pedigree from a great-grandparent of the respective hybrid (DGGP) for 9% of hybrids, and one class (2% of hybrids) with an inbred ranked into first place that was directly related by pedigree to the great-great-grandparent of that hybrid. Inbreds that ranked in second position were related to the respective parents of the hybrid as follows: Thirty-one (57% of hybrids) were a grandparent of the respective hybrid, 11 (20%) were classed as DGP, 7 (13%) were DGGP, 1 (2%) was class DGGGP, and 4 (7%) were a great-grandparent (GGP) of the respective hybrid. Over all hybrids, two of the four grandparents ranked into first and second positions for 23 (43% of hybrids); three grandparents ranked into the first three positions for 5 (9% of hybrids). There were no instances where all four grandparents ranked into the first four positions. Thirty hybrids had a grandparent ranked into first position using $P = 0.99$. The number of grandparents ranked into the top five positions was 98 (compared to 108 when $P = 0.50$). The number of grandparents ranking into the top two positions was 55 (compared to 71 when $P = 0.50$). The mean probability of a grandparent that ranked into the first two positions was 0.9288 (SD = 0.1454) when $P = 0.50$ and 0.9980 (SD = 0.0104) when $P = 0.99$.

DISCUSSION

The prevalent use of paternity indices demonstrates that it is advantageous to have explicit probabilities of ancestry to distinguish among different pedigrees. Molecular marker profiles are rapidly becoming more extensive and cost effective to generate. Features that would advance the statistical analysis of molecular marker data to provide explicit probabilities of ancestry include the ability to calculate probabilities of ancestry where there is no *a priori* information as to the identity of one (usually the maternal) parent and robustness in the face of laboratory error.

Maize inbred lines and hybrids provide a very exacting set of materials for evaluating the discriminatory abilities of molecular data and statistical procedures that are employed to interpret those data. Hundreds of maize inbred lines of known pedigree together encompass a great diversity and complexity of pedigree relationships. Some inbred lines can be very highly related and genetically similar due to their derivation from common parentage including from parents that are themselves highly related. Consequently, relationship categories such as "sister" or "parent" when applied to maize inbreds usually refer to closer degrees of pedigree relationship and, thus, of germplasm and molecular marker profile similarity than those of the equivalently named classes of relationship for animal species. Most maize hybrids that are widely used in the United States today are constructed from pairs of inbred lines that are unrelated by pedigree, each inbred parent having been bred from a separate "pool" of germplasm. Various degrees of relatedness are possible between hybrids according to the pedigree relationships among their constituent inbred parents.

Using $P = 0.99$ in the algorithm is more specific for identifying parents than using $P = 0.50$. However, $P = 0.99$ is less robust for identifying other relatives, such as grandparents. When the algorithm was run at $P = 0.50$ there were 6 hybrids for which one parent did not rank among the top two most probable genotypes. For the remaining 48 hybrids the correct parents were identified even in circumstances where other candidate inbreds included not only full-sister lines bred from related parents but also inbreds even more closely related to the true parent by virtue of being backcross conversions of the inbred parent of the hybrid. For each of the 6 hybrids where a nonparent ranked above a true parent, that higher ranked inbred was always either a sister or progeny of the outranked true parent. The range of pedigree relationships as expressed by the Malécot coefficient of relatedness (Malécot 1948) that was encompassed by pairs of true parents and more highly ranked inbred relatives of the true parents was from 0.8390 to 0.9680. A coefficient of 0.8390 approximates a relationship between inbred A and A′ where

inbred A' has been bred from a cross of inbreds A and B with between one and two additional backcrosses of the parental inbred A. A Malécot coefficient of relationship of 0.9680 closely approximates a relationship between inbreds A and A" where four additional backcrosses of parental inbred A follow the initial cross of inbreds A and B.

Running the algorithm at $P = 0.99$ in comparison to $P = 0.50$ raises the probability of ancestry for the parents while diminishing the probabilities for the third and lower ranking candidate inbred lines. Use of the algorithm at $P = 0.99$ increased both the percentage of hybrids with both parents ranked in the first two positions (from 89 to 96%) and the percentage of parental inbreds that were ranked first and second (from 94 to 98%). Two hybrids (3914 and X0915A) did not have both parents ranked first and second when the algorithm was run at $P = 0.99$. For both of these hybrids the nonparental inbred that outranked the true parent was itself a product by pedigree from the true parent that had been created by an additional four backcrosses of that parent; the Malécot coefficient of relationship between the parent of the hybrid and the inbred that outranked that parent for these two hybrids was 0.9636.

Robustness was tested by evaluating the effects of using data from different numbers of loci and by simulating additional levels of missing and misscored data up to combined levels of 25% error beyond that which was provided by the laboratory. From our experience, error rates of 5 to 10% can occur in SSR profiling of maize due chiefly to the combined effects of residual heterozygosity among seed lots and by deficiencies in the scoring of heterozygotes in hybrids. The additional levels of simulated error, therefore, include values (up to ~35% total error) that are well outside of our experience. For five hybrids that were examined, increasing the number of loci from 50 to 100 (with no additional missing or misscored data) did reduce the number of instances where inbreds that were not parents of a hybrid outranked the true parent from four to one. Nonetheless, all of these more highly ranked inbreds, although they were not themselves the true parents of the respective hybrid, were either direct progeny or full sisters of the true parent (Table 2). Consequently, if such degrees of error can be tolerated in respect of pedigrees for inbreds that are identified as parents of hybrids, then SSR data from 50 loci of equivalent discrimination ability are sufficient. Use of data from 50 loci also evidenced robustness in the face of up to 10% additional levels of either missing or misscored data; no degradation in the ability to identify a parent was apparent up to the level of 10% additional error except for 10% additional missing and misscored alleles for one hybrid (3525; Table 3). However, use of 100 loci increased the proportion of true parents that were correctly identified from 53% (for 50 loci) to 71% (mean correct parents over all levels of error; Table 3). Use of data from 195 loci provided greater resiliency against additional levels of error. However, use of data from 195 loci was unable to provide resiliency against the negative effects of adding combined levels (at 25%) of both missing and misscored data (Table 3). At the 25% level of additional poor data integrity, inbreds that were not related to the true parent of the hybrid outranked the true parent for four of the five hybrids. Levels of missing or misscored data should, therefore, be kept below 15–20% (assuming a level of 5–10% error in the data we analyzed prior to simulating additional error).

We have previously examined the pedigrees of inbreds that are ranked into the first two positions when the true parents are removed from the list of candidate inbred lines. Usually, direct progeny or full sisters of the true parents then rank most highly (data not presented). We therefore examined the rankings of inbreds with respect to their ranking and probability of inclusion in the ancestry of each hybrid after the removal, not only of the true parents, but also of the progeny of the true parents and any full sisters of the true parents. In these circumstances the grandparents of the hybrids are ranked predominantly into top positions. Using $P = 0.50$, a grandparent ranked into first position for 76% hybrids and into second position for 57% hybrids; with $P = 0.99$ a grandparent ranked into first place in 56% of hybrids. At $P = 0.50$ two grandparents ranked into first and second positions for 43% hybrids and into the first three positions for an additional 9% hybrids. Most of the remaining inbreds that ranked into the top two positions were progeny of the grandparent. A total of 108 grandparents ranked into the top five positions when $P = 0.50$; 93 ranked into these positions when $P = 0.99$. Seventy-one grandparents ranked into the top two positions when $P = 0.50$; 55 grandparents ranked into these positions when $P = 0.99$. The mean probability of a grandparent in the top two positions was 0.9288 (SD 0.1454) when $P = 0.50$ and 0.9980 (SD 0.0104) when $P = 0.99$. Our algorithm was written to identify pairs of ancestors; alternative algorithms could be tailored to identify all grandparents once parents had been identified and removed from the list of candidate inbreds.

We have demonstrated the capability and robustness of an algorithm that can be used to show probability of parentage in circumstances where the *a priori* pedigree identity of neither parent is known. Exclusions are taken into account, thereby allowing parentage to be shown even when the two parents are not represented in the database of molecular profiles that are examined. Heterozygous candidate parents can be accommodated. The number of loci that is necessary to provide a reliable basis of determining pedigree is dependent upon the degree of relatedness among parents and nonparents and upon the discriminatory ability of the marker system

in the species of interest. Using $P = 0.99$ compared to $P = 0.50$ preferentially identified more true parents and with a greater difference of probability to third placed nonparents. If there is reasonable assurance that the parents are among the candidate list of inbreds, then $P = 0.99$ should be used; if greater robustness is required, then $P = 0.50$ should be used.

Applications of our algorithm include the identification of pedigrees among individuals of plant or animal species where molecular profile datasets exist that can be interpreted in terms of segregating alleles at individual marker loci and that provide a sufficient power of discrimination. Capabilities to generate large datasets of suitable molecular profile data are already available and are increasing rapidly with the advent of single nucleotide polymorphisms. One further application of our algorithm is to assist in the protection of intellectual property that is obtained on plant varieties or upon specific dams or sires of animals through the determination of pedigrees.

LITERATURE CITED

Alderson, G. W., H. L. Gibbs and S. G. Sealy, 1999 Parentage and kinship studies in an obligate brood parasitic bird, the brown-headed cowbird (*Molothrus ater*), using microsatellite DNA markers. J. Hered. 90: 182–190.

Bein, G., B. Driller, M. Schurmann, P. M. Schneider and H. Kirchner, 1998 Pseudo-exclusion from paternity due to maternal uniparental disomy 16. Int. J. Leg. Med. 111: 328–330.

Berry, D. A., 1991 Inferences using DNA profiling in forensic identification and paternity cases (with discussion). Stat. Sci. 6: 175–205.

Berry, D. A., 1996 *Statistics: A Bayesian Perspective*. Duxbury Press, Belmont, CA.

Berry, D. A., and S. Geisser, 1986 Inferences in cases of disputed paternity, pp. 353–382 in *Statistics and the Law*, edited by M. H. DeGroot, S. E. Fienberg and J. K. Kadane. Wiley Publishing, New York.

Bockel, B., P. Nurnberg and M. Krawczak, 1992 Likelihoods of multilocus DNA fingerprints in extended families. Am. J. Hum. Genet. 51: 554–561.

Bowers, J. E. and C. P. Meredith, 1997 The parentage of a classic wine grape, Cabernet Sauvignon. Nat. Genet. 16: 84–87.

Chakraborty, R., T. R. Meagher and P. E. Smouse, 1988 Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. Genetics 118: 527–536.

Chakraborty, R., L. Jin and Y. Zhong, 1994 Paternity evaluation in cases lacking a mother and nondetectable alleles. Int. J. Leg. Med. 107: 127–131.

Devlin, B., K. Roeder and N. C. Ellstrand, 1988 Fractional paternity assignment: theoretical development and comparison to other methods. Theor. Appl. Genet. 76: 369–380.

Elder, J. K., and E. M. Southern, 1987 Computer-aided analysis of one dimensional restriction fragment gels, pp. 165–172 in *Nucleic Acid and Protein Sequence Analysis—A Practical Approach*, edited by M. J. Bishop and C. J. Rawlings. IRL Press, Oxford.

Ellstrand, N. C., 1984 Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. Am. Nat. 123: 819–828.

Glotz, K. and G. Pfeiffer, 1998 Assignment of individuals to populations using microsatellites. J. Anim. Breed. Genet. 115: 53–61.

Gunn, P. R., K. Trueman, P. Stapleton and D. B. Krabbowski, 1997 DNA analysis in disputed parentage: the occurrence of two apparently false exclusions of paternity, both at short tandem repeat (STR) loci, in the one child. Electrophoresis 18: 1650–1652.

Hamrick, J. L., and A. Schnabel, 1985 Understanding the genetic structure of plant populations: some old problems and a new approach, pp. 50–70 in *Population Genetics in Forestry*, edited by H. R. Gregorius. Springer-Verlag, Heidelberg, Germany.

Helminen, P., V. Johnsson, C. Ehnholm and L. Peltonen, 1991 Proving paternity of children with deceased fathers. Hum. Genet. 87: 657–660.

Lang, J. W., R. K. Aggarwal, K. C. Mukundar and L. Singh, 1993 Individualization and estimation of relatedness in crocodilians by DNA fingerprinting with a Bkm-derived probe. Mol. Gen Genet. 238(1–2): 49–58.

Malécot, G., 1948 *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris.

Marshall, T. C., J. Slate, L. E. B. Kruuk and J. M. Pemberton, 1998 Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639–655.

Meagher, T. R., 1986 Analysis of paternity within a natural population of *Chamaelirium luteum* (L). Identification of most-likely male parents. Am. Nat. 128: 199–215.

Meagher, T. R., and E. Thompson, 1986 The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theor. Popul. Biol. 29: 87–106.

Miller, P. S., 1975 Selective breeding programs for rare alleles: examples from the Przewalski's horse and California Condor pedigrees. Conserv. Biol. 9: 1262–1273.

Primmer, C. R., M. T. Koskinen and J. Piironen, 2000 The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. Proc. R. Soc. Lond. B Biol. Sci. 267: 1699–1704.

Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA 94: 9197–9201.

Seeg, K. M., H. Steinkellner, J. Glossl, S. Kampfer and F. Regner, 1998 Reconstruction of a grapevine pedigree by microsatellite analysis. Theor. Appl. Genet. 97: 227–231.

Smith, J. S. C., E. C. L. Chin, H. Shu, O. S. Smith, S. J. Wall *et al.*, 1997 An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor. Appl. Genet. 95: 163–173.

Smouse, P. E., and T. R. Meagher, 1994 Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) gray (Liliaceae). Genetics 136: 313–322.

Taylor, A. C., A. Horsup, C. N. Johnson, P. Sunnucks and B. Sherwin, 1997 Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat *Lasiorhinus krefftii*. Mol. Ecol. 6: 9–19.

Thompson, E., and T. R. Meagher, 1987 Parental and sib likelihoods in genealogy reconstruction. Biometrics 43: 585–600.

Vankan, D. M., and M. J. Faddy, 1999 Estimations of the efficacy and reliability of paternity assignments from DNA microsatellite analysis of multiple-sire matings. Anim. Genet. 30: 355–361.

Wetton, J. H., D. T. Parkin and R. E. Carter, 1992 The use of genetic markers for parentage analysis in *Passer domesticus* (house sparrows). Heredity 69: 243–254.

White, E., J. Hunter, C. Dubretz, R. Brost, A. Bratton *et al.*, 2000 Microsatellite markers for individual tree genotyping: application in forest crime prosecutions. J. Chem. Technol. Biotechnol. 75: 923–926.

Communicating editor: Z.-B. Zeng

What is an "Essentially Derived Variety"?

The concept of essentially derived variety was introduced into the 1991 Act of the UPOV Convention in order to avoid plagiarism through mutation, multiple back-crossing and to fill the gap between Plant Breeder's Rights and patents, gap which was becoming important due to the development of the use of patented genetic traits in genetic engineering.

An essentially derived variety is a variety which is distinct and predominantly derived from a protected initial variety, while retaining the essential characteristics of that initial variety.

As indicated as an example in the UPOV Convention, essentially derived varieties may be obtained by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, back-crossing, or transformation by genetic engineering.

The commercialization of an essentially derived variety needs the authorization of the owner of the rights vested in the initial variety.

The concept of essentially derived variety does not at all abolish the Breeder's Exemption, as free access to protected plant varieties for breeding purposes is maintained. It is not a threat to biodiversity. On the contrary, it favors biodiversity, encouraging breeders developing and marketing original varieties.

Appendix C
Serial No. 09/490,884

UPOV Publication No. 644(E), Section 1

www.upov.org

INTERNATIONAL CONVENTION

FOR THE

PROTECTION OF NEW VARIETIES OF PLANTS

of December 2, 1961, as revised
at Geneva on November 10, 1972,
on October 23, 1978, and
on March 19, 1991

adopted by the Diplomatic Conference
on March 19, 1991

reproduced from UPOV Publication No. 438(E)
issue No. 63 of "Plant Variety Protection"

1991 Act of the Convention

Article 12
Examination of the Application

Any decision to grant a breeder's right shall require an examination for compliance with the conditions under Articles 5 to 9. In the course of the examination, the authority may grow the variety or carry out other necessary tests, cause the growing of the variety or the carrying out of other necessary tests, or take into account the results of growing tests or other trials which have already been carried out. For the purposes of examination, the authority may require the breeder to furnish all the necessary information, documents or material.

Article 13
Provisional Protection

Each Contracting Party shall provide measures designed to safeguard the interests of the breeder during the period between the filing or the publication of the application for the grant of a breeder's right and the grant of that right. Such measures shall have the effect that the holder of a breeder's right shall at least be entitled to equitable remuneration from any person who, during the said period, has carried out acts which, once the right is granted, require the breeder's authorization as provided in Article 14. A Contracting Party may provide that the said measures shall only take effect in relation to persons whom the breeder has notified of the filing of the application.

CHAPTER V
THE RIGHTS OF THE BREEDER

Article 14
Scope of the Breeder's Right

(1) [Acts in respect of the propagating material] (a) Subject to Articles 15 and 16, the following acts in respect of the propagating material of the protected variety shall require the authorization of the breeder:

(i) production or reproduction (multiplication),

(ii) conditioning for the purpose of propagation,

(iii) offering for sale,

(iv) selling or other marketing,

(v) exporting,

(vi) importing,

(vii) stocking for any of the purposes mentioned in (i) to (vi), above.

(b) The breeder may make his authorization subject to conditions and limitations.

(2) [Acts in respect of the harvested material] Subject to Articles 15 and 16, the acts referred to in items (i) to (vii) of paragraph (1)(a) in respect of harvested material, including entire plants and parts of plants, obtained through the unauthorized use of propagating material of the protected variety shall require the authorization of the breeder, unless the breeder has had reasonable opportunity to exercise his right in relation to the said propagating material.

(3) [Acts in respect of certain products] Each Contracting Party may provide that, subject to Articles 15 and 16, the acts referred to in items (i) to (vii) of paragraph (1)(a) in respect of products made directly from harvested material of the protected variety falling within the provisions of paragraph (2) through the unauthorized use of the said harvested material shall require the authorization of the breeder, unless the breeder has had reasonable opportunity to exercise his right in relation to the said harvested material.

(4) [Possible additional acts] Each Contracting Party may provide that, subject to Articles 15 and 16, acts other than those referred to in items (i) to (vii) of paragraph (1)(a) shall also require the authorization of the breeder.

(5) [Essentially derived and certain other varieties] (a) The provisions of paragraphs (1) to (4) shall also apply in relation to

(i) varieties which are essentially derived from the protected variety, where the protected variety is not itself an essentially derived variety,

(ii) varieties which are not clearly distinguishable in accordance with Article 7 from the protected variety and

(iii) varieties whose production requires the repeated use of the protected variety.

(b) For the purposes of subparagraph (a)(i), a variety shall be deemed to be essentially derived from another variety ("the initial variety") when

(i) it is predominantly derived from the initial variety, or from a variety that is itself predominantly derived from the initial variety, while retaining the expression of the essential characteristics that result from the genotype or combination of genotypes of the initial variety,

(ii) it is clearly distinguishable from the initial variety and

(iii) except for the differences which result from the act of derivation, it conforms to the initial variety in the expression of the essential characteristics that result from the genotype or combination of genotypes of the initial variety.

(c) Essentially derived varieties may be obtained for example by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, backcrossing, or transformation by genetic engineering.

Article 15
Exceptions to the Breeder's Right

(1) [Compulsory exceptions] The breeder's right shall not extend to

(i) acts done privately and for non-commercial purposes,

(ii) acts done for experimental purposes and

(iii) acts done for the purpose of breeding other varieties, and, except where the provisions of Article 14(5) apply, acts referred to in Article 14(1) to (4) in respect of such other varieties.

(2) [Optional exception] Notwithstanding Article 14, each Contracting Party may, within reasonable limits and subject to the safeguarding of the legitimate interests of the breeder, restrict the breeder's right in relation to any variety in order to permit farmers to use for propagating purposes, on their own holdings, the product of the harvest which they have obtained by planting,